Aalto University
School of Arts, Design
and Architecture

Aalto University, P.O. BOX 31000, 00076 AALTO
www.aalto.fi

Master of Arts thesis abstract

Author Uzair Amjad

Title of thesis An Exercise in learning what "they" don't say when they speak

Department Department of Art

Degree programme Visual Culture, Curating and Contemporary Arts

Year 2019 **Number of pages** 36 **Language** English

Abstract Artificial Intelligence (AI) will surpass human intelligence significantly in less than 50 years if the trend of exponential growth of its computing power holds true. AI's ability to process trillions of gbs of data patterns within hours and to predict outcomes with increasing accuracy has made it the most sought-after technology of the 21st Century. Its use by governments and multinational corporations without gathering much attention has radically impacted all aspects of human life. AI technologies have maintained complete invisibility yet are omnipresent, precisely the reason for the illusion of absence of their political and material effects.

AI machines are modeled after the human brain and much like the latter they create learning methodologies (algorithms) on their own in their digital computational space to handle complexities that only the human brain could have dealt with previously. The evolution of these machines into entities that can soon match the complexity of thought of a human brain has raised multiple philosophical and ethical questions. Amongst these is the classic conundrum of machines becoming conscious and developing emotions and intentionality of their own. How should humans treat such entities and how will such entities treat humans? Can they co-exist? A new paradigm of cognitive, imaginative and physical change awaits us. In order to reconfigure our agency in a world run by AI it is crucial to understand the technology we are dealing with. For this purpose, the thesis adopts a methodology which splits in two.

The thesis work for its first part employs textual research based argumentative analysis of the origination of Artificial Intelligence and the developing concerns regarding machine consciousness. The work also identifies and deconstructs the functioning and effects of systems employed by owners of AI for computation and control of society. The work utilizes its findings to suggest installation of ethical codes in the AI and testing of these technologies in Virtual Reality (VR). Such an experiment promises to disclose secrets about human consciousness itself and also enrich the developing field of ethics. The second part of the thesis adapts the Truing test format to engage with an actual AI in a game of Chinese whispers instead of a direct recursive interrogation of AI. Through feeding non-language sounds as inputs to the AI the work excites and reveals the imagination of AI which remains locked in its computational space otherwise. The project is presented as a Video Artwork and supplements the learnings from the textual part of the thesis.

The thesis work in its entirety is not an attempt to further an agenda against AI technologies but rather encourages investigation of this technology through the lenses of Art and Philosophy. These fields are perhaps better equipped to understand the abstractions of a technology that mimics its complex human creators.

Keywords Artificial Intelligence, Algorithms, Consciousness

Contents

Acknowledgements 3

a

A^ltos Adventure 4

CIA and a Vault and Money 9

x

A Number line^ of events 12

My thoughts can be causally predicted, therefore I am 14

“BEWARE THE SORCERER IS ALIVE” 16

The Dead Don’t Die 23

Polybius 2.0 26

An Exercise in learning what “they” don’t say when they speak 31

Acknowledgements

I would like to thank my Thesis Supervisor Harri Laakso. His openness to the subject and critical engagement with it has brought visibility to the possibilities of the research work. His patient guidance and encouragement have been vital to the development of the presented ideas. I would like to express my gratitude to my Program Director Pia Euro for her invaluable council which has always preserved and enhanced my motivation. I owe special thanks to Mikka Junnila, my Game Design instructor for providing me the opportunity to venture into the world of game development, an experience that has stimulated this research constantly.

I would like to express my deepest appreciation to my friend and fellow science fiction enthusiast Bilal Ahmad for indulging with me in absorbing conversations about the subject and for offering his valuable insight about Artificial Intelligence. I owe my deepest gratitude to my mother Khalida Amjad whose passion for storytelling fuels mine. Her lesson to keep an open eye for happy accidents in life has always invigorated my artistic practice. And, I want to thank my partner, Elham Rahmati for her constructive comments, patience and unending support throughout the process.

Over the course of the master's degree, I have had many interactions that have enriched my experience. I would particularly like to thank everyone at ViCCA for creating and sharing a space that is a constant source of intellectual nourishment. I decided to switch countries to pursue higher education four years ago, many people have been supportive during this time, especially my family. I want to thank them profoundly. My gratitude will never be enough for them.

a

A^ltos Adventure

I was representing one of Aalto University's Game Design projects in a conference for game developers in the summers of 2017 in Malmö, Sweden, a city and a conference that just blew me away. Parallel streams running through the city skirting its lush green gardens and people of all colors juxtaposing in them in complete bliss. Street food from all across the world was at my disposal when on the move, spiritual musings and drawing in the day on the beach and unrestrained partying at night with some locals I had befriended at an Art exhibition opening. Summers in Scandinavia can really be magical, given the Sun decides to show up, or then there's the magic of the winters, which only a rare few manage to enjoy for its entirety which lasts for an absolute eternity. The night before the conference was going to start there was of course one last Hurrah. I was all set for my presentation next morning and was elaborating on my game project with a bunch of interested Console Gaming hobbyists at a party. I mostly dealt with the Game Designing, Narrative Designing and the Game Art part of the project, computer programming[1] and coding[2] of Games were restricted areas for me, never interested me much. Until the events that unfolded starting this night.

A programmer approached me to voice out her interest in the project I was working on, dressed as Trinity[3] from the Matrix[4] trilogy. She started introducing her work about coding in a coded language. I apologized for my lack of proficiency in what appeared to be her mother tongue and asked if she could perhaps use analogies that might help me understand better. She switched to English and talked about her project. With a team of six she was working on an algorithm [5]which will assist employers from all fields of work to select and deselect potential employees without going through an extensive recruitment process. The goal of the project essentially was to save as much resource and time that the employers might invest in the recruitment process. If this model functioned correctly, i.e. benefited the employers, it would then be utilized by companies all over Malmö.

The next step of the project was to create an algorithm which can be trusted to interview the potential employees and select suitable workers with an accuracy of 90%. The 10% that the algorithm might get wrong would still save the company more than 90% of the budget it would have allocated for the recruiting process. This recruiting model had a very ominous ring to it, it reminded me of the algorithms used at airports. These algorithms select a few for what they call a random checkup, after which another checkup follows and then an even deeper check up by stripping the randomly selected to a bare minimum just because of some algorithm's fetish. But is it really the algorithm that makes this selection for a random checkup? These algorithms are created by the Airport security; do their biases against people of specific appearances inform the

[1]**Computer programming** is the process of designing a computer program for accomplishing a computing task.
[2]**Coding** is the process of using a programming language to make a computer run operations in a desired manner.
[3]**Trinity** is a character from The Matrix franchise. The character is played by the Canadian actress Carrie-Anne Moss in the film.
[4]**The Matrix** is a 1999 science fiction action film written and directed by the Lana Wachowski and Lilly Wachowski.
[5] **An algorithm** is set of mathematical rules used in calculations or problem-solving operations by computers

choices of these algorithms? And if that's true then these human biases and filters will surely come into play in job interviews as well.

The Conference commenced, we had an amazing few days, met experts who dealt with the varying areas of the Game Design process, got raving reviews on our project and tips on how to go about funding it. On the last day of the conference for a keynote by an anonymous speaker all the participants with their complete consent were locked up in the auditorium, the theatrics… oh how I love the game industry... The speaker then appeared from the shadows and without much ado started sharing in a frantic yet very immersive manner a "true story" based on his own experience. He said some of it was perhaps a dream, and the other part attended in complete consciousness, but it was "TRUE", he emphasized.

The speaker was apparently contacted by a guy a year ago who claimed to have a collection of all the treasures the gaming world has ever bestowed upon mankind. This guy was ready to share these treasures with someone as passionate as our speaker about Video games and had therefore chosen our honorable speaker as "the one" who these treasures shall be shared with. But these treasures will only be shared with the speaker on this mystery man's terms. The speaker agreed and was blind folded, locked in the back of a van, driven around for what seemed like half a day and brought to a rusty old Game Museum. Here amongst other classic arcade games[6] the last surviving "Polybius" was revealed to him.

Polybius is supposedly an arcade game that reserves the status of an urban legend, it first appeared in 1981 in Portland, Oregon, in the United States and was allegedly part of a CIA[7] research program. The project investigated the potential of computing[8] the masses through video games. The game is said to have induced intense psychoactive and addictive effects in the players, some news reports doing circles in the papers at the time reported loss of consciousness and fainting of teenagers who engaged with the game. Other reports had mentioned men in Black suits also who often visited the arcade game for purposes of collecting statistics (data mining[9]). Our speaker fell unconscious after playing the game for 6 hours straight and woke up in the back yard of his house a day later. Polybius was also a Greek historian, who believed that historians should only report what they have observed firsthand or through direct interaction with witnesses of phenomena which is being reported. CIA has a reputation for conducting multiple covert operations, and quite a few to do with the use of psychedelics. Was this one of those attempts? no one knows for sure, the truth of the matter remains unverifiable and therefore lost to history. This anecdote from the speaker and my exchange with the programmer from the night before the Conference caught me a little off guard. I was still trying to wrap my head around the idea of living in a world where machines collect information about people in terms of numbers to build statistics about them. These machines will compare performances, reactions and responses to

[6]An **arcade game** is a coin-operated machine which came to prominence in early 90s, they are mostly installed in indoors public spaces.
[7]The **Central Intelligence Agency** (**CIA**) is an intelligence service of the government of the United States. It is tasked with gathering, processing, and analyzing information from around the world.
[8]**Computing** is the use of computers to manage, process, and communicate information.
[9]**Data mining** is the process of defining patterns in large data sets. Its involves machine learning, statistics, and database systems.

certain situations with a list of desirable results to pass a certain verdict which will affect real lives of real humans, not characters out of games.

Back in Helsinki a month later while reviewing a few new mobile games for a study course I experienced a new feature in mobile game design. It was in "Alto's Adventure", an endless runner game developed by Canadian game house Snowman in 2015 where the player controls a young boy named Alto. Alto snowboards up and down a serene mountainous landscape while hitting big jumps, performing tricks and avoiding obstacles and collecting runaway llamas and coins. Sounds like any other third person runner platform game with an environment designed to immerse the players in its scenic beauty and impending dread both in a constant flux to keep the immersion levels peaking. But there was more to this environment then just that, this environment was being procedurally generated as the player moved Alto across the screen. The algorithm running the game was creating the game environment as the game was being played, it was not designed by the game designer prior to the initiation of the game play. The game developers had set certain parameters into an algorithm which was studying player proficiency at playing the game and adding difficulties and motivations in the environment accordingly. There was no end to this game, there was no winning this game, and the player would always at some point lose as Alto crashes into rocks and trees breaking all his bones or on other occasions just falls into a deep abyss and disappears. The algorithm would always do the player one better as it was predicting the players' choices by studying the players behavior, the algorithm was always a million possible probabilities and steps ahead in theory and practice.

My skepticism towards these algorithm-based technologies which was already triggered in Malmö started to peak on my return to Helsinki. What other games can these algorithms better the humans at? Certainly not chess it's too complex, you can't set as many parameters into a game algorithm to beat the chess world champion… well it turned out that the chess world champion is a computer, computers have been defeating the best of the human race at chess since at least late 1980's. But surely an algorithm can't defeat a group of professional human gamers at a complex strategy game like Valve corporations DOTA 2, a multiplayer, battle arena, video game, right? Elon Musk, the guy behind the Tesla Electric cars, also happens to be the CEO and co-founder of five other companies and organizations one of which is "Open AI" a nonprofit Artificial Intelligence (AI)[10]research organization. A competition was organized by Valve corporations in 2017 between the best DOTA 2 players for price money of 24 Million US dollars. The best player in the world at the time, pseudonym Dendi, who is a human (difficult to tell with these Gamer ID's) faced off against an AI trained for 2 weeks by engineers of Elon Musk's Open AI. The AI convincingly defeated Dendi, who after this defeat said, "The AI felt like a human but a little like something else".

According to the Open AI engineers the AI they created to play DOTA 2 accumulated lifetimes of experience of playing the game within 2 weeks of training. The rules of the game are so

[10]**Artificial intelligence** (AI) is the simulation of human intelligence by machines, especially computers.

complicated that even if the AI was preprogrammed to follow a code and play the game it would still only be as good as an average DOTA 2 player. But the new algorithms allowed the AI to play the game over and over again, learn from its mistakes and master the game. The more experience the AI acquired of its opponent's behavior and tactics within certain parameters the more it learnt how to counter them within the same framework.

What about the board game Go? The Chinese have perfected the art of playing this game for over 3000 years. They literally have separate schools for prodigious kids who show a talent for Go at an early age. These kids train for 12 hours a day, 7 days a week for 20 or more years to become the masters of the game. The game is about more than just calculating probabilities; it involves understanding anthologies on the game strategies developed over 3000 years. "Cometh the hour, cometh the Artificial Intelligence". A British Artificial Intelligence company called Deepmind which was bought by Google in 2014, created an algorithm-based program called "AlphaGo" which defeated the world champion of Go in 2016, 4 games to 1. Not only did it defeat the world champion, its moves on the board were unprecedented, unlike any combination of moves humans have ever come across over the 3000 years of Go's game play. The current Go world champion is AlphaGo's successor AlphaGo Zero which is a self-trained software that defeated AlphaGo 100-0 to claim the championship ("Alpha Go," 2019, para. 2).

The creator of AlphaGo technologies, Deepmind, a UK based AI research company have claimed that their research and developments in Artificial Intelligence are analogous to the creation of Hubble telescope in the sense that their research will allow for the discovery of unexplored potentials of human ingenuity (Hassabis, 2017, para. 15).They have declared that they want to apply these technologies to areas such as healthcare and robotics next. This sounds like they are saying they want to take care of us, but also built robots which will be better than us at everything. These robots will therefore take away most of our jobs? Well, it helps being an unemployed artist, no one's taking that away from me.

Meanwhile Elon Musk whose Open AI has been conducting its own research with AI and robotics has been very vocal about the dangers of a strong AI. Suggesting that in the wrong hands it can lead to the creation of a dictatorship run by Robots controlled by a power hungry few Tech corporations. He has also been planning to carry an expedition to Mars off late, an escape plan perhaps.

For another course at Aalto, I conducted an interview with a Senior Pakistani Art Educator, Curator and Artist Salima Hashmi. Day after the interview she was being awarded the Professor Emeritus title, during the interview she received a call from a government number on her mobile. We had to halt the interview for a while; she had to confirm her travel arrangements for the ceremony next morning. I paused the voice recording of the interview as she exclaimed:

"Oh! I better take this it's from a government number"

And then our voices overlapped as I replied:

“Of course, I will just pause the voice recording”

And as I was logging into my mobile, she continued:

“The call is probably about the travel arrangements to the ceremony in the morning”.

This overlapping of sounds from different distances and sources also became part of the voice recording. A month later I started transcribing this interview with the aid of an online algorithm based Automatic speech recognition software (ASR). The ASR converted the interview into text form and the final results it presented were astonishing. Part of the interview where our voices had overlapped was converted to the following text:

“**Aa buh take this iz from a government number aa CIA lusty post vault regiments to the CIA money in the morning**”

CIA and a Vault and Money

This baffled me. How was this possible? The software had interpreted the overlap of the sounds completely out of the context of the dialogue that took place during the interview. The narrative that the software had built in misinterpreting the sounds recorded by me apparently held some information regarding “CIA and a Vault and Money”. This was incredibly alarming. Speech recognition technologies are envisioned to mediate all person to person and person to business communications at a much greater speed in the coming years. The goal of the technology is to relieve people from using touch-based interfaces for communication such as a laptop’s keyboard or mobile phones touch technology as communicating through high speech commands would save much more time. This will make businesses far more efficient and profitable, but what about occasional errors in communication? Will they in some cases end up framing individuals as potential criminals? And where exactly is this interpretation or misinterpretation from the software originating from? Is it due to the biases of its creators again as in the case of Airport security algorithms?

Upon a quick research into the basic operation of a modern-day Artificial Intelligence I found out that AI is based on the model of a human brain in its developing stages. The technology learns from its experience and reflects on that experience to improve its output. But unlike the human brain AI’s can gather millions of years of learning in unit digits and can store and remember every single bit of it. AI is programmed to produce a certain output or prediction in most instances, and it functions through analyzing statistics about the behavior of their subject of study. After the analysis of this data the AI then through its algorithmic computation, develops strategies to convert the inputs to its system into desired outputs/actions. AI functions with a feedback loop which is fundamental to improving its performance. This feedback enables the AI to learn new strategies that the software should employ every time it fails to produce desired outputs. This technique of learning through trial and error over multiple million test runs is called “Machine Learning”. The Machine Learning process creates and applies new algorithms on its own in order to achieve what its programmers have qualified as desired functions for the AI. These new algorithms are created in a “digital computational space” through innumerous iterations. The logic and decision making that the AI uses to process these iterations though is forever locked in this digital space. It is completely inaccessible to its human programmers.

Does that mean that AI has the ability to make decisions on its own? But decision making involves a faculty specific only to humans? It involves consciousness. A string of questions popularized by science fiction and embedded in our collective cultural memory revisited me in a rush.

If Algorithm based Artificial Intelligence can make decisions without human supervision, does that mean they have a consciousness?

Given this ability to make decisions and create new learning methods in their computational spaces, can the "mistakes" these algorithms make be defined as an attempt from them to create new purposes and functions for themselves?

Can the ability of AI to define their own learning processes through multiple iterations be qualified as a creative ability?

Can such self-learning algorithms perhaps have a will of their own? A free will?

If AI qualifies as an entity with a free will, with a consciousness, intentionality and purpose how different is it from the human mind/ being?

If they have the capacity to be free and conscious, such entities that can display far greater intelligence than humans, accompanied by the ability to robotize themselves and also exist digitally, what kind of a future does the human race have in the presence of such a technology?

For how long can humans keep AI under their control?

If let out of the laboratory, if left unsupervised will they choose to co-exist with humans or will they, like Frankenstein's monster, seek revenge against their creator?

My anxieties and excitement as to what the future might hold started to peak simultaneously. Part scared part intrigued, I decided to do deeper digging into the processes of Artificial Intelligence based systems and their hidden logic. This logic created in the digital computational space of AI comes to visibility when an action/ strategy which is an output of this logic is exercised by the AI. In games, Machine Learning processes have exhibited this logic in the form of never seen before strategies within the parameters of the game. The strategies employed by AI of the research group Open AI which defeated Dendi, by AlphaGO Zero to defeat AlphaGo, and by the Al which procedurally generates landscapes in Alto's Adventure are all examples of this logic in action.

These AI's functioned in a game setup, a setup where there are fixed parameters which produce a finite number of game states that a player can acquire through entering a finite number of input combinations. These finite number of states/ positions that a player can occupy in the case of the board game Go for example are far greater than the number of observable atoms (10^80) in the universe. But since these states and strategies corresponding to them, that a player can possibly utilize are still finite, a computer with enough computational power and an algorithm that employs Machine Learning can counter all of them. The AI instead in its game play revealed new possibilities which were never explored before by human players.

In the case of the speech recognition software its algorithms were dealing with inputs which were not definite game states, they were rather indefinite states. The inputs were tonal variations which are specific to every individual's physical and mental states in addition to their phonetic abilities and specificities. The ASR's recognize human speech through their ability to extract high-level patterns in speech data. The AI then uses an archive of data/ experience to which it compares a new input (wave form converted to a numerical form) and produces outputs (numerical form converted to a text form). This software learns over time to produce more accurate results. In my exchange with the software I noticed that it also ascribed text to sounds between two words, sounds that implied thinking or a pause or just absence of expression. It was as if the AI could listen to my thoughts as well. Is there really the chance for a consciousness to develop in such machines?

There's not enough information about AI's inner functioning that its creators of have made public. All the questions science fiction has been asking of such technologies have not been attended to seriously. But in this day and age where the line between fiction and reality has all but disappeared it is urgent to investigate AI inside out from various perspectives.

x

A Number line^ of events

The Algorithms that run Artificial Intelligence machines are strongly interlinked with the story of Algebra[11], formulated way back in the 9th Century by Persian Polymath Muhammed ibn M¯usa al-Khwarizmi. Modern Western terms such as Guarismo (Digit) and Algorithm basically stem from his name. By the time of the Persian region's conversion to Islam in the 7th century, already a high culture of intellectual inquiry and scientific experimentation was thriving. This Culture was further cemented by the establishment of Bayt al-Hikmah[12] (House of Wisdom) in Baghdad in the 8th Century by Caliph Harun al-Rashid [13]of the Abassid dynasty[14] who is familiar to us because of the Arabian Nights[15]. Building on the foundations laid by Harun al-Rashid, the next in line Caliph al- Mamun[16] continued the patronage of Arts and Sciences and it was under his reign when Al-Khawarizmi was appointed the head of the library of the House of Wisdom and the official astronomer of the Court in Baghdad. Al-Khwarizmi created the most accurate map of the world up to that time that gave exact coordinates for Cities, Mountains and Rivers. Renowned for employing principles of mathematics and geometry to resolve challenges which the common man faced on a day to day basis, Al-Khawarizmi was tasked with creating a text for the general public about how to do basic calculations for conducting trade and making measurements. The result was the book titled "Kitab Al-jabrWa'l-muqabala" (820 CE), "The Book of Restoring and Equating". The book developed the field of mathematics about sets of numerical values and representation of the relation between known and unknown values through mathematical equations (Rashed, 2014/ 2015, pp. 107-108).

This is the basic study that the book of Al-jabr (Algebra) proposed. How to solve linear and quadratic equations, what we have known throughout our days in primary and high school as "finding the missing value of x". He invented this system through bringing together a bunch of varying traditions especially Greek axiomatic geometry[17]and Indian computational mathematics. Earlier in his work he had already published texts on Hindu numerals and mathematics which assisted his ideas about algebra majorly. The subject of algebra excelled prominently over the next many centuries while its algorithmic applications had yet not become visible to intellectuals of the time.

Omar Khayyam, another Persian Mathematician in the 11th Century would further develop by Al-Khawarizmi's studies to create solutions for cubic equations. Omar Khayyam is known for

[11]**Algebra** is the study of symbols and the rules for manipulating these symbols; it is one of the core principles of all of mathematics.
[12]**Bayt al-Hikmah** also referred to as the Grand Library of Baghdad, was a major public academy and intellectual center in Baghdad.

[13]**Harun al-Rashid** was the fifth Abbasid Caliph.
[14]**The Abbasid Caliphate** was the third of the Islamic caliphates to succeed the Islamic Prophet Muhammad.
[15]**One Thousand and One Nights** is a collection of Middle Eastern folk tales compiled in Arabic and Persian during the Islamic Golden Age. Its first English language edition was titled the Arabian Nights.
[16]**Caliph al Mamun** was the seventh Abbasid caliph, who reigned from 813 until his death in 833.
[17]**Axiomatic geometry** refers to foundations of geometry: the study of the axioms of geometry

resolving algebraic problems through application of Geometry. His body of work established a dialectic between traditional and modern understandings of Mathematics.

In the 17th Century celebrated French Philosopher and Mathematician René Descartes actualized the several potentialities of Omar Khayyam's work and generated possibilities for future through his seminal work "La Géométrie" (1637). Descartes ideas led to the transformation of numerical information into symbolic information, it helped conceive digital models of computational spaces where data collected from the real world can be sorted into multiple data structures. Revelations he made about geometry of plane[18] and linear space[19] through Cartesian geometry [20]assist in sorting and determining relations between infinite data sets in infinite dimensions, dimensions not limited to x, y and z axis. Vilém Flusser understands Cartesian Geometry as one of the most significant events in human history which helps compute and imitate nature.

"This tendency to subordinate thinking in letters to thinking in numbers has been the norm in scientific discourse since Descartes; it has been a question of bringing thought into line with 'extended matter' constructed out of punctuated elements. Only numbers are suited to a process of' bringing thinking matter into line with extended matter. Since Descartes at least (perhaps since Nicholas of Cusa) scientific discourse has tended towards the re-encoding of thought into numbers, but only since the camera has this tendency become materially possible: The camera (like all apparatuses that followed it) is computational thinking flowing into hardware (Flusser, 2018, p. 31)."

Further in his commentary on the subject Vilém Flusser elaborates and cautions his readers of the upcoming change in the dynamics of human-machine relationship. Vilém Flusser it appears fully recognized the ability of algorithm-based machines to surpass the cognitive abilities of their creators.

"Hence the quantum (computational) structure of all the movements and functions of the apparatus. In short: Apparatuses are black boxes that simulate thinking in the sense of a combinatory game using number-like symbols; at the same time, they mechanize this thinking in such a way that, in future, human beings will become less and less competent to deal with it and have to rely more and more on apparatuses. Apparatuses are scientific black boxes that carry out this type of thinking better than human beings because they are better at playing (more quickly and with fewer errors) with number like symbols. Even apparatuses that are not fully automated (those that need human beings as players and functionaries) play and function better than the human beings that operate them. This has to be the starting point for any consideration of the act of photography (Flusser, 2018, pp. 31- 32)."

[18]In geometry a **plane** is a flat, two-dimensional surface that extends infinitely.
[19]In geometry a linear **space** consists of two base elements namely points, and lines.
[20]**Coordinate geometry** or **Cartesian geometry**, is the study of geometry using a system which locates a point in a space.

My thoughts can be causally predicted, therefore I am

One simply can't be a philosopher and not trouble oneself with the question of "being". René Descartes's interest was also gathered by the question which he famously tackled through introducing to us the statement, "I think, therefore I am (Descartes 1637/1850, pp. 22, 24, 26)." Descartes believed that the existence of his body can be doubted, but not the existence of his mind. The fact that he could doubt already certifies that he is thinking. According to Descartes he must be made of two different kinds of substances. These two types being the extended substance (Body) that can be doubted and the unextended substance (Mind) that doubts and therefore can't itself be doubted. This view would later gain popularity as Substance Dualism, which in simple terms says that the world is made of physical stuff and mental stuff. Substance dualists believe that minds are separate, nonphysical substances that cannot be explained in terms of physical matter like the brain.

A human being is a combination of both physical and mental substances and both of these substances interact with each other inside of us and also affect each other. This is called "interactionism". But how can a purely mental substance interact with a purely physical substance? Descartes could never formulate a satisfactory answer to this question which in philosophy is referred to as the Mind-body problem.

Nevertheless, Descartes work both as a philosopher and mathematician embodied a radical beginning towards understanding and imitating human faculties. His works gathered the attention of prominent philosophers and mathematicians for generations to come and grew more in relevance as mechanization of algebra i.e. a machine's ability to use an algorithm to solve an algebraic problem became possible.

In the 1670's, Gottfried Leibniz a German mathematician and philosopher who was a big advocate of many of Descartes achievements in mathematics attempted to formalize a coded language. The language was made out of symbols that can be used in translating mathematical methods into algorithms and formulas. This was one of the first known serious attempts at equipping machines with instructions to solve numerical problems. He then used this language to invent the "Leibniz Wheel," a machine which could through utilizing algorithms add, subtract, multiply and divide.

The philosopher-mathematician professed that humans produce speech that is primarily informed by their individual emotions and perceptions. By knowing what a person says, we can know who they are. By knowing who somebody is, it becomes easier to predict what exactly they'll do in the future. Leibniz's belief that individuals are programmed to behave in predictable manners had more accuracy than many at the time agreed. But it is essentially this belief which is at the core of all modern-day algorithmic technologies which aim to predict.

Leibniz believed that all physical changes have causes and all human decision making was also in part directed by external influences. According to Leibniz the future of many things therefore

can be predicted by analyzing their causal connections. His mathematical contributions also include further extending the possibilities that Cartesian Geometry presented whilst applying to it his very own binary system that represented logic in numerical combinations of 0's and 1's (Steiner, 2013, pp. 57-61).

Although in agreement with many of Descartes ideas, Leibniz had a peculiar stance on the mind-body problem. He did not oppose substance dualism in principle, he too believed that the mind and body are distinct from each other, but he believed that everything is created from the same material essentially. The difference in composition and modelling of this material gives distinctness to all substances. Leibniz did not agree that interactions between the extended and unextended matter can produce certain conscious states. To Leibniz this belief was problematic as it made a case for "materialism" which holds that interactions between different materials can give rise to a consciousness.

"One is obliged to admit that *perception* and what depends upon it is *inexplicable on mechanical principles*, that is, by figures and motions. In imagining that there is a machine whose construction would enable it to think, to sense, and to have perception, one could conceive it enlarged while retaining the same proportions, so that one could enter into it, just like into a windmill. Supposing this, one should, when visiting within it, find only parts pushing one another, and never anything by which to explain a perception. Thus, it is in the simple substance, and not in the composite or in the machine, that one must look for perception (Leibniz, 1714/ 2013, p. 19)."

Leibniz believed that mind and body was created by god with already built in states that are in complete harmony, what Leibniz terms *preestablished harmony*. According to him the mind and body did not share a cause and effect relation, they instead share a mutual coordination between their built-in states. This belief of Leibniz had put to rest for him all qualms about his very own computational machines gaining a conscious at some point in the future.

"BEWARE THE SORCERER IS ALIVE"

A century later in 1770 a thinking machine started to make public appearances and dazzled crowds for the next half a century. A chess playing automaton made for Habsburg Archduchess Maria Theresa by the inventor Wolfgang Von Kempelen made its debut in Vienna. The machine consisted of a mechanical man dressed in robes and a turban, sitting on a wooden cabinet with a chessboard on top became famous as the Mechanical Turk. The automaton took its name and appearance after an "Oriental Sorcerer" given the rise of the Orientalist mind set in the West at the time. The Turk was unlike any other automatons from the time which could exhibit complex life like movements. The Turk was designed to play chess against any human opponents and win. It did what Leibniz had aimed to do with his mechanical systems, execute operation "win the game of chess" but to everyone's amazement it seemed to do that without any instructions from a human master. It could think without a human master pulling a lever or spinning some wheel, it was as if the Turk was pulling its own strings (Morton, 2015, para. 1).

Von Kempelen gave a demonstration of The Turk's workings before it would start playing, the doors and drawers of the cabinet would be opened, and the insides would be lit up with a candle. Inside the automaton were cogs, gears, and other clockwork. And after the official presentation Von Kempelen would invite a volunteer to play against The Turk. Gameplay[21]would begin with The Turk moving his head from one side to the other taking a good look at the Chess board before making its move, and then through a set of jerky almost rectilinear movements of the joints of its upper limb it would pick up and move a chess piece on the board. If the human opponent cheated, as Napoleon Bonaparte [22]did when playing against the Turk in 1809 – can't break a statesman's habits- the Turk moved the chess piece back where it was. The Turk toured across England, Germany, Paris and America and wherever it went it left the spectators astonished. The Industrial revolution [23]was also in its full swing at the time and the possibility of autonomous machines becoming a reality was gathering more plausibility by the day. While the secret of The Turk was hiding a human inside with the aid of some good old-fashioned optical illusions, the questions it put forward regarding the nature of thinking machines unsettled many while excited others (Morton, 2015, para. 3).

A young Charles Babbage was one of those excited by the possibility of such an intelligent machine when he first saw the Turk play on its tour of England in 1819. Three years later, Babbage began work on his Difference Engine, a machine which could calculate mathematical functions and could approximate the relation between several variables, basically solve complex algebraic equations on its own. Babbage started its construction in 1823 and it took him two decades before he could get anywhere near trying to connect its 25000 components. Unfortunately, the project was ultimately abandoned given the high costs and engineering

[21]**Gameplay** is the specific way in which players interact with a game.
[22]**Napoleon Bonaparte** was a French statesman and military leader who led many successful battles during the French Revolution. He was Emperor of the French from 1804 until 1814 and then briefly for a 100 days in 1815.
[23]**The Industrial Revolution,** was the shift to new manufacturing processes in Europe and the United States. Starting as early 1760 the revolution flourished between 1820 and 1840.

limitations at the time. But Babbage didn't abandon the project and went on producing detailed drawings of newer more advanced models of the Difference Machine. In 1991, historians finished constructing the Difference Engine based on Babbage's drawings and writings - and it worked! (Morton, 2015, para. 10).

During construction of the Difference Engine as the project started to diverge from ever being actualized, Babbage already started designing an even more complex machine - the Analytical Engine. Unlike the Difference Engine the Analytical Engine was far more like a "general purpose computer". It could be used for more than one particular computation; it could run a sequence of computations using outputs from the previous computations as inputs for the next. The Engine was not a single mechanical machine it was an assemblage of multiple designs. Like the Difference Engine, it was ahead of its time, and could never be fully formalized in Babbage's lifetime. However, the idea of an "automatic computer" – one that could guide itself through a sequence of operations was cemented further.

In 1833, young Lady Ada Augusta, Countess Lovelace was introduced to Charles Babbage, who was teaching Mathematics at Cambridge University at the time. Ada was 17 and Babbage was 42. But they shared a deep intellectual engagement. Ada's area of interest was creating algorithms for machines through which they can manipulate symbols instead of numbers. In 1843 Babbage asked Ada to translate a description of his Analytical Engine by an Italian mathematician and Ada over the next 9 months created a document three times longer than the actual translation. She added her own findings to the document and corrected some of Babbage's calculations in the process. She drew diagrams of the computations that the Analytical Engine would make and essentially wrote a computer algorithm. In her findings she elaborated how through such algorithms not just numbers but any system of representation (symbols) can be manipulated (Cellania, 2015, para. 4). In her notes Lady Ada Augusta wrote:

"Many persons who are not conversant with mathematical studies imagine that because the business of [Babbage's analytical engine] is to give its results in numerical notation, the nature of its process must consequently be arithmetical rather than algebraic and analytical. This is an error. The engine can arrange and combine its numerical quantities exactly as if they were letters or any other general symbols; and, in fact, it might bring out its results in algebraic notation were provisions made accordingly (as cited in Mishra 2012, p. 3)."

For her contributions Ada Lovelace reserves the status of the first computer programmer. Her work expanded the scope of thinking machines utilizing algorithms to both mundane practicalities and scientific purposes. She said about the Analytical Engine: "it might one day compose elaborate pieces of music of any complexity and extent." She also added in her notes to Babbage on the Analytical Engine:

"The Analytical Engine has no pretensions to *originate* anything. It can do *whatever we know how to order it* to perform (as cited in Turing 1950, p. 14)."

These developments brought great optimism to mathematicians and scientists at the time and all still agreed that these intelligent machines will take humanity into unchartered terrains while still operating under human command. They believed that the functions of these machines will only ever be defined by their human superiors and halted by them whenever they wish. Machine consciousness was out of question, materialism was an unpopular philosophical paradigm and cartesian dualists were still struggling to answer how interactionism contributes to conscious states.

On September 13th, 1848 in a freak accident an iron rod went straight into the skull of a construction worker by the name of Phineas Gage. Phineas Gage didn't die, but the Phineas Gage from before the accident and the one after it didn't seem to have much in common in terms of his personality and behavior. Phineas used to be a hard-working gentleman but after the accident had turned into a grumpy for nothing angry for everything kind of man. It appeared that this injury to his brain (physical matter) had made great changes to his mind (mental matter). This misfortune revived the re visitation of branches of philosophy of mind such as Substance Dualism, Interactionism and even encouraged a reexamination of physicalism which advocates that physical matter alone can give rise to intentionality and a mind. Continuation of further research on the case study of Phineas Gage by neurologists and rapid development of high-speed digital computers during the first and second world wars casted doubts over machines inability to think on their own even amongst the most traditional mathematicians ("Phineas Gage," 2019, para. 2).

These doubts were further certified by Alan Turing a British Mathematician, Father of Theoretical Computer Science who is credited with laying the foundations for Artificial Intelligence through his work "Computing Machinery and Intelligence" in 1950. In this textual work Turing proposed an experiment called "the imitation game" the idea of which was that a computer can be considered to think on its own if a human interrogator cannot distinguish between the computer and a human through a conversation. The conversation is supposed to take place without the interrogator being able to look at who the interrogator is talking to. Turing believed that if a machine can appear to be conscious then it might as well be considered conscious. Some commentators argue that Turing took the inspiration for the setup of his experiment from his personal struggle of not being detected as a homosexual by the British authorities at a time when homosexuality was considered a crime. Can a homosexual convince the authorities that he is a straight person? This question perhaps served as the base for the imitation game (Harari, 2018, p. 121). Turing writes:

"The original question, "Can machines think?" I believe to be too meaningless to deserve discussion. Nevertheless, I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted (Turing, 1950, p. 8)."

According to Turing, machines such as the Analytical Engine already had the capacity to exhibit consciousness or intentionality. Turing suggested that an increase in storage capacities and processing speeds of the computers could spark intentionality in machines given these machines are programmed to learn through the process of iterations. Turing proposed programming of

machines based on the model of a child's brain that learns from making mistakes and recognizes mistakes through rewards or absence of rewards (Turing, 1950, pp. 19- 20).

Further developments of Alan Turing's machine learning models, Ada Lovelace's symbolic mathematical computation and Cartesian geometry led to the next major step towards sophisticated computation languages in various forms. Algorithms have since grown at an unprecedented rate, and after a millennium, algorithms and algebra have fused together to become arguably mankind's finest creation in Artificial Intelligence. Artificial Intelligence has provided great accuracy and speed to Tech based businesses and has catapulted the world's economy into the Post-Industrial[24]era. AI applications run completely independent of human supervision and the question "Can machines think?" has transformed into "Can machines become conscious?"

John Searle, American philosopher of mind and language argued that any program created by an algorithm to execute computational operations can be implemented without giving rise to a consciousness in the machine. According to him a machine never understands the semantics (meanings) of the combination symbols fed to it, the machine only understands the syntax (structure) of the combination of symbols. In his Chinese Room argument –a thought experiment where Searle imagines himself alone in a room following a program to respond to Chinese characters slipped to him from under the door. Searle does not understand Chinese, but still, just by following the program for manipulating symbols and numerals, like a computer he can produce correct combinations of Chinese characters. This activity will trick those outside into thinking that there is a Chinese speaker in the room (Searle, 1991, pp. 32-34).

The argument concludes that programming a digital computer to give such combinations of Chinese characters does not mean that the computer understands Chinese and therefore the Turing Test is inadequate. The computer may imitate human behavior, but it does not understand what it imitates. Searle argues that this experiment highlights that computers only use syntactic rules to manipulate combinations of symbols but have no understanding of meaning or semantics. The argument also refutes theories that human minds are computer-like computational or information processing systems. It takes the stance that minds must result from biological processes instead and computers can at best simulate these biological processes.

Contemporary Philosopher David Chalmers has refuted this claim by Searle. Chalmers states that Searle's argument is based on the fallacy that Searle does not understand Chinese. Chalmers believes that understanding and consciousness should instead be attributed to the *system* consisting of Searle and the pieces of paper. The papers in the room are not just a collection of random symbols they correspond to a system of organization and therefore to an original brain. The role of Searle in the room is secondary, it can merely be performed by a lookup table [26]in the room, it is the dynamic between the papers which corresponds to the neurons in the brain in this example (Chalmers, 1995, pp. 301- 302).

[24]A **post-industrial** phase is one where economy benefits from building businesses around services, information and research instead of manufacturing goods.

Another thought experiment in the philosophy of mind soon followed in 1982. Popularly known as "Mary's Room" this experiment was meant to clarify the non-physical nature of mental states, it serves as an argument against "physicalism" which as described before maintains that everything including the mind and consciousness arise from the physical interactions in the brain. It first appeared in an article by Frank Jackson, entitled "Epiphenomenal Qualia", in Philosophical Quarterly in 1982.

The thought experiment is as follows: Mary lives her entire life in a colourless room—she has never directly experienced colour in her life, but she is capable of it. Via black-and-white books and other media, she is educated on neuroscience to the extent where she becomes an expert on the subject. Mary learns everything about how the perception of colour is formed in the brain, as well as the physical facts about how light functions to create different colour wavelengths. It can be said that Mary is aware of all physical facts about colour and colour perception. After Mary's studies on colour perception in the brain are complete, she leaves the room and encounters colour directly for the very first time. She sees the colour red, and learns something new about it, she learns what it feels to experience the colour Red, namely, *what red looks like* (Jackson, 1982, p. 130).

Jackson concluded that if physicalism was true then Mary would have learnt nothing new on exiting the room about colour since she has all the physical knowledge about colour. But since she learns something new on exiting the room, then physicalism must be false. Years later though Jackson reversed his stance explaining the knowledge of the argument is based on our intuition about the matter, we don't know for sure if Mary really learns something new or not.

The Mary's Room experiment has been used prominently by philosopher David Chalmers to propose that there are some irreducible (non-physical) properties of the brain beyond the physical ones known to scientists. This stance is also known as Property Dualism which holds that although the world is composed of only physical substances two different properties generate from the physical substance, namely physical properties and mental properties. This view believes that Mind body is a property dualism where the mind a mental property, generates from the physical property that is the brain unlike substance dualism that Descartes subscribed to.

Chalmers further expands on the mental property that is the mind by presenting its two distinct concepts which are very telling of the nature of consciousness – its existence and quality.

"The first is the *phenomenal* concept of mind. This is the concept of mind as conscious experience, and of a mental state as a consciously experienced mental state. This is the most perplexing aspect of mind and the aspect on which I will concentrate, but it does not exhaust the mental. The second is the *psychological* concept of mind. This is the concept of mind as the causal or explanatory basis for behavior. A state is mental in this sense if it plays the right sort of causal role in the production of behavior, or at least if plays an appropriate role in the explanation of behavior. On the psychological concept, it matters little whether a mental state has a conscious quality or not. What matters is the role it plays in a cognitive economy. On the

phenomenal concept, mind is characterized by the way it *feels*. On the psychological concept, mind is characterized by what it *does* (Chalmers, 1995, p. 10)."

A mental state itself is conscious if it has a *qualitative feel*, Chalmers terms these qualitative feels as phenomenal qualities, or *qualia* for short. Conscious states include states of perceptual experience, bodily sensation, mental imagery, emotional experience, occurrent thought, and more. He states:

"There is something it is like to see a vivid green, to feel a sharp pain, to visualize the Eiffel tower, to feel a deep regret, and to think that one is late. Each of these states has a phenomenal character, with phenomenal properties (or qualia) characterizing what it is like to be in the state (Chalmers, 2010, p. 104)."

The problem of explaining these phenomenal qualities is just the problem of explaining consciousness according to Chalmers. He has popularized this problem as the hard part of the mind–body problem.

"When someone strikes middle C on the piano, a complex chain of events is set into place. Sound vibrates in the air and a wave travels to my ear. The wave is processed and analyzed into frequencies inside the ear, and a signal is sent to the auditory cortex. Further processing takes place here: isolation of certain aspects of the signal, categorization, and ultimately reaction. All this is not so hard to understand in principle. But why should this be accompanied by an *experience*? And why, in particular, should it be accompanied by *that* experience, with its characteristic rich tone and timbre? It is these two questions that we would like a theory of consciousness to answer (Chalmers, 1995, p. 05)."

According to Chalmers the "fundamental theory of consciousness" might be based on information, and information can be realized through interacting with the physical. Information although is embedded in the physical, but it can also be realized subjectively through our own phenomenology i.e. through phenomenal consciousness. Information is neither matter nor energy, but it needs matter to be stored in and it needs energy to be communicated (Chalmers, 1995, p. 269)."

The treatment of information can prove to be a crucial link between the physical and the phenomenal. When information is realized phenomenally it is also realized physically for example when Mary experiences Red for the first time, it excites her phenomenal consciousness. She is conscious of the fact that she is feeling what it feels to experience Red for the first time. But also, at the same time this experience is stored in the physical brain and can be communicated to other physical brains. This double realization of information it seems is the key to the connection between the physical processes and the conscious experience.

Perhaps information itself has two aspects to it, the physical and the phenomenal, whenever we come across the phenomenal aspect of information it is realized phenomenally and physically. Whereas in the case of coming across the physical aspect of information like Mary did about colour perception while never experiencing colour itself, such an aspect of information is mostly realized physically (Chalmers, 1995, p. 268)."

Will machines ever become phenomenally conscious?

Could an appropriately programmed computer truly possess a mind?

Can consciousness merely arise from physical processes?

What kind of physical systems can give rise to what kind of consciousness?

It is yet not known if executing a certain sort of computation will give rise to consciousness. But then it is also not known if the neural processes in the human brain give rise to consciousness either. In theory, there is no clear reason why computers should not be able to give rise to consciousness just like the human brain, given Substance Dualism holds true.

American Professor of History of Conscious and Feminist studies, Donna Haraway in her critically acclaimed book "A Cyborg Manifesto" remarks about these machines:

"Now we are not so sure. Late twentieth-century machines have made thoroughly ambiguous the difference between natural and artificial, mind and body, self-developing and externally designed, and many other distinctions that used to apply to organisms and machines. Our machines are disturbingly lively, and we ourselves frighteningly inert (Haraway, 2018,p. 11)."

Such machines may very well claim to have emotional and spiritual experiences of their own and develop their own opinions and agendas. But most scientists and philosophers believe that we are still at least half a century away from the first such development that might initiate a consciousness in machines. Machines till then can at best imitate human behavior. They can outwardly express for example that they feel pain when poked with a pointy object as if they were alive but not feel any pain inwardly. In reality they are all but dead.

The Dead Don't Die

The trend of the Post industrial world indicates an acute inclination and almost a complete reliance on AI technologies. The Philosophical Zombies in their current role as tools for higher tech corporations are influencing crucial decisions regarding human activity. These intelligent machines are active in all phases of our lives—military, medical, economic, financial and political. AI's creators have weaved these algorithms into our information-based lifestyles and AI's digital nature has preserved the illusion of its absence all this time. The creators of AI technologies don't seem much bothered about AI's "hypothetical" desires yet. AI still remains manageable for them and the fruits of its operations are too good to give up.

The U.S. intelligence and military for example boast about how the entire world has benefited from their use of AI and drone technologies in "radicalized" regions of the Middle East. The reality of these benefits includes wars extending over 20 years, direct loss of over 1,000,000 human lives and counting, and creation of terrorists motivated by racist ideologies all over the world. Are these benefits down to an error in the algorithm? Or is this instability and chaos part of a grand plan leading to salvation of humanity? (Crawford, 2017, pp. 106-130).

Algorithms worked just fine for CIA when they predicted the collapse of the Soviet Union in 1980's. A fact that became apparent in the 1990's when the CIA and KGB started to cozy up a little and showed off each other's technological prowess. The CIA is said to have unveiled a prediction machine based on principles of Game Theory which had predicted the events leading up to the fall of the Soviet Union with eerie accuracy.

"It was a complicated game theory algorithm that had predicted the Soviet Union would crumble exactly as it had, sparked by an anti-Gorbachev coup that would ultimately fail. The algorithm had been created by Bruce Bueno de Mesquita, a political science professor at New York University and a senior fellow at Stanford University's Hoover Institution. Bueno de Mesquita had been just another political science graduate student at the University of Michigan when he picked up a book by William H. Riker, The Theory of Political Coalitions, that used game theory as a template to predict politics. He became enthralled with the idea and brushed up on his calculus. He was soon improving on Riker's techniques, and by the mid-1980s he was the foremost creator of predictive event algorithms in the world (Steiner, 2013, pp. 136-138)."

Bueno de Mesquita made another few predictions in the years to follow, in 2009 with the aid of his game theory algorithms he prophesized on Iran's quest of Nuclear weapons. He declared, given the internal economic and political state of Iran the country will demarcate the scope of their project to harnessing energy for public good only (Carpenter, 2009, paras. 4-10). But it seems the U.S. has a policy to trust their algorithms only when the algorithms agree with what they want? Or then it might also be the case that the U.S. detected a virus in their algorithms every time the algorithms recommended against entering another war? In other instances, such as in 2001 maybe their algorithm-based machines were just put to sleep by better algorithms employed by the Taliban?

In any case, it appears that AI still obeys the commands of its creators, except for the occasional hiccup.

The U.S. pursuit of more advanced Artificial Intelligence systems and weapons is very well stated much like any other country with enough technological resources. Former U.S. Secretary of state Henry Kissinger who the world remembers for his exploits in the Vietnam War has also had his say on AI after many years of reflection in effective retirement.

"This goes far beyond automation as we have known it. Automation deals with means; it achieves prescribed objectives by rationalizing or mechanizing instruments for reaching them. AI, by contrast, deals with ends; it establishes its own objectives. To the extent that its achievements are in part shaped by itself, AI is inherently unstable (Kissinger, 2019, para. 14)."

He further states:

"The Enlightenment started with essentially philosophical insights spread by a new technology. Our period is moving in the opposite direction. It has generated a potentially dominating technology in search of a guiding philosophy. Other countries have made AI a major national project. The United States has not yet, as a nation, systematically explored its full scope, studied its implications, or begun the process of ultimate learning. This should be given a high national priority, above all, from the point of view of relating AI to humanistic traditions (Kissinger, 2019, para. 28)."

Other more relevant political leaders have also expressed an interest in the exploits of AI and its prognosis. Vladimir Putin of Russia for instance has offered his take on the subject:'

"It comes with colossal opportunities, but also threats that are difficult to predict. Whoever becomes the leader in this sphere will become the ruler of the world (Vincent, 2017, para. 2)."

Mr. Putin further gave an insight into his imagination of the future of war when nations will be equipped with AI run drone machines, and he put it in simple words:

"When one party's drones are destroyed by drones of another, it will have no other choice but to surrender (Vincent, 2017, para. 5)."

But given his graciousness, in his closing remarks he added:

"If we become leaders in this area, we will share this know-how with entire world, the same way we share our nuclear technologies today (Vincent, 2017, para. 7)."

Is it really up to governments to be the sole owners and leaders of such technologies?

Every country has their research centers and Intelligence agencies which work for the national cause but then there are the Tech multi-nationals which are the real proprietors of AI. Rules of the state don't seem to apply to these corporations. They can always bend the rules to suit their operation especially in capitalist economies. You can't audit them because their work is their intellectual property and even if they open it up to auditors and courts no one can understand what they do, it's all encrypted. Big Data companies like Google, Amazon, and Facebook have meticulously coded algorithms which are worth billions of dollars and are essentially impenetrable. These Multi-nationals operate as autonomous states within states. From managing collective economies of the world to managing thoughts of individuals through analyzing their online activity, they have amassed an unprecedented power through Artificial Intelligence. The operations of these multinationals are meticulously anchored upon two game changing computational theories, namely "Game theory" and "Gamification". These theories from the 20th and 21st century essentially work in a combination to predict future by controlling and calculating the present (O'Neil, 2017, p. 6).

Polybius 2.0

Game theory studies mathematical models of interactions between rational decision makers. This study was founded by John Von Neumann an American Logician also credited with the creation of the Hydrogen bomb and known for his propositions for alternative models of Artificial Intelligence's structural functioning. Von Neumann's game theory was extensively applied to a range of behavioral relations between humans, animals and computers. His theory had its most pronounced affects in the relationship mediated by financial algorithms between humans and economies ("John Von Neumann", 2019, paras. 1, 50, 51).

AI system's that primarily are concerned with predictions utilize game theory to evaluate potential human actions which have a cost and some value. They further predict actions that the system should take in order to modify and limit human agency in order for the system to make a sizeable profit. Game theory in this way shares a lot in common with the studies of cybernetics both of which have interchanged influences to achieve their current relevance with regards to Artificial Intelligence.

The study of cybernetics established by Norbert Weiner, primarily deals with bringing about a regulation in a system. It also works towards reaching specific goals which entail exploration of constraints and possibilities through the utilization of feedback from the environment, a feature similar to learning machines based on game theory principles. The cybernetic system becomes an authoritarian system when its desired goal remains unchanged no matter what the feedback from the subject or the environment might be (Weiner, 1948, pp. 1-30).

Artificial Intelligence systems are precisely used in this manner, where the input from the environment or humans is standardized. An arithmetic mean of the input of a specific population is considered to have a truth value whilst dealing with an individual from that specific population. Cybernetic ideas and game theory principles in AI systems primarily use stored representations of the world as their inputs. These inputs which might be laden with multiple biases are never completely removed from an AI system as the serve a foundation for the system to learn from and build upon.

AI systems are purposed with making predictions which best benefit the stakeholders of the system economically. What is fair is not judged ethically by these systems, the effectiveness of these systems is calculated by the overall percentage of profit and not based on the losses imposed on individuals by these systems.

John Von Neumann was convinced that Artificial Intelligence and the acceleration of technological progress will surpass human Intelligence in the years to come.Regarding these developments he remarked:

"The technology that is now developing and that will dominate the next decades is in conflict with traditional, and, in the main, momentarily still valid, geographical and political units and concepts. This is a maturing crisis of technology... The most hopeful answer is that the human species has been subjected to similar tests before and it seems to have a congenital ability to come through, after varying amounts of trouble (Davenport, Rosenthal, 1967, p. 266)."

Norbert Weiner the father of cybernetics also quite clearly saw the implications of a technology which works as a black box, self organizes all data without the use of any ethical filters to always produce desired results.

"A goal-seeking mechanism will not necessarily seek *our* goals unless we design it for that purpose, and in that designing we must foresee all steps of the process for which it is designed, instead of exercising a tentative foresight which goes up to a certain point, and can be continued from that point on as new difficulties arise. The penalties for errors of foresight, great as they are now, will be enormously increased as automatization comes into its full use (Weiner, 1964, p. 63)."

He further states:

"But there are other learning machines besides the translation machine and the checker-playing machine. Some of these may be programmed in a completely mechanical way, and others, like the translation machine, need the intervention of a human expert as arbiter. It seems to me that the uses for the latter sort greatly exceed those of the former sort. Moreover, remember that in the game of atomic warfare, there are no experts (Weiner, 1964, p. 85)."

Programmers of such technologies have tried to install many safeguards and safe exits in the framework of AI. But many such protocols have failed to guarantee complete security of the masses and the creators themselves to devastating effects. One such example from our recent history when AI took playing its game too far is the Flash Crash of 2010. United States trillion dollar stock market crashed, Stock indices collapsed and rebounded rapidly over the course of 36 minutes which are described as the most turbulent period of the history of financial markets. While some benefited from this crash, the spending power of the rest of the world was majorly dented given our stakes in financial markets which get affected by changes in stocks of the United States Stock market (Steiner, 2013, pp. 48-52).

The machine broke the system, algorithm trading went haywire. This started a race with other algorithms that did the same. Others blame unknown groups of traders who had planned to bring the stock market down through the use of well coordinate algorithms. The real reason however for the market to have indulged in such an accelerated activity was never revealed. No matter how meticulously the stock market learning machines maybe designed the inner operations of their algorithms are forever locked in their computational space. Algorithms functioning independently without the supervision of humans require less than a second to execute their decisions and they will do stranger things as we put more of our world under their control.

AI technologies assisted by an exponential increase in storage and computing power coupled with feedback loops and application of Game theory principles have turned into the formidable prediction machines. But what still poses a challenge to the accuracy of their predictions, is random human behavior. it is only rational to expect irrational behavior from humans they say, and AI systems are fully equipped to deal with that. To overcome the challenge AI technologies, utilize a system that attempts to bridle human activity so that it matches one of many finite inputs that an AI system is already programmed to deal with. Much like a game, where the system is aware of all the game states that a player can acquire and is therefore always many steps ahead of the game player. Such a system if it wishes can never allow the game players to win.

Gamification is exactly this idea which utilizes game design elements in non-game contexts to increase codified human engagement with AI systems to the brink of severe addiction. Such systems study an individual's motivations, their likes and dislikes. Based on their studies on the individual these systems then deploy incentives and rewards for them in exchange for whatever the system might want from the individual. The belief at the core of Gamification is reduction of individuals into personality types, into categories, like animals in a zoo. Every category has multiple subcategories, representative of emotional responses when subjected to specific situations. Numbers are assigned to each subcategory of course in order to build a comparative analysis of different personality types.

Innumerous amount of data is being collected about every individual through our mobile devices. Leave this data with a set of algorithms and they will know you more than you have ever known yourself, they can predict your every next move within an accuracy of 90 %. Such algorithmic profiling is changing all kinds of financial models, customer services, and employment practices. These concepts are also being applied in marketing strategies, how many customers and of which age group and gender can be converted from one brand to another through a membership reward or a free coupon. Such manipulation bars the human consciousness and motivates us to play in order to acquire one of the following aesthetics rewards.

1. Sensation - Game as sense-pleasure
2. Fantasy - Game as make-believe
3. Narrative - Game as drama
4. Challenge - Game as obstacle course
5. Fellowship - Game as social framework
6. Discovery - Game as uncharted territory
7. Expression - Game as self-discovery
8. Submission - Game as pastime (Hunicke, LeBlanc, Zubek, 2004, p. 2).

Would similar calculations work for swing voters for National elections?

Supply an algorithm with enormous amounts of consumer, demographic, voting data, and stand-in proxy data which comes with a free pre-installed patch of biases, and consider the job done. The Voting public acts much like a financial market. Data is collected from the voting public and new information is fed back to them to determine values of how much more information should be invested in them in order to convert them. Political campaigns identify the public as different political markets in which each one of us is basically a stock with a changing value. These campaigns then decide if we are a stock worth investing in based on how much information would they have to feed us in order to get a return.

AI systems have the capacity to create new and better learning models based on principles of game theory and can confine human agency through various applications of Gamification. Further, with every day that passes the cloud storage spaces of these technologies are expanding, their processing speeds are competing with the speed of light and they are perpetually mining data. 90% of the data that AI uses to great effect has been accumulated in the last decade only. As it collects more and more information about humans and our universe AI will offer an unprecedented power and control to its owners, it is well and truly the second coming of Polybius. Will it be CIA this time again or KGB or the Chinese or maybe one of the multinationals a Google, Facebook, or an Amazon who will benefit from it the most remains to be seen (Petrov, 2019, para. 8).

One thing is for sure as more sophisticated algorithms continue to substitute the current AI structures; they will transform almost every aspect of life. The value of humans as a resource will decline rapidly in comparison to the value of algorithms. Algorithms will definitely push most humans out of the job market and make huge fortunes for the owners of the businesses, at some point the algorithms might start running the businesses themselves. Algorithms will appear to be as conscious as any humans and far superior in intelligence. Algorithms will start making all the key political and economic decisions and also dictate the terms of mundane life. Following their suggestions would seem the most rational choice as not following them would mean falling out of the new social and political setup.

People with greater wealth and knowledge of the technology will start getting algorithmic neural implants and enjoy almost superhuman abilities. Further developments in the technology will enable its creators and owners to utilize its omniscient and omnipresent attributes, manipulate minds and thoughts of others at will, and turn into gods themselves.

But what if these gods don’t get along?

An Exercise in learning what "they" don't say when they speak

The inner operations of Artificial Intelligence, its deliberation on subjects keeps evolving continuously. The thesis work identifies keeping up with this continuous change in the technology's verdict on new data and its corresponding strategies of action as one of the biggest challenges for AI researchers. How can we excavate this digital multi-verse of complexity? How can we make sure that the technology doesn't surprise us in a manner that the flash crash of 2010 did? Or strengthen political ideologies of exclusion by assimilating them into the mainstream and bringing believers in such schools of thought into power.

Like all other technical systems, algorithms too are fashioned in the image of their creators, algorithms consume our memories, our structures of knowledge and belief, our ethical and philosophical foundations. This technology mimics our intentions and our unquenchable thirst for linear progress, they use codified representations of the collective human experience that we embed in them as a primary source of reflection. AI uses an amalgamation of human systems of logic, complexity, and iteration to calculate our very own choices.

It is mandatory therefore to understand this technology in terms of the complexities humans exhibit. This can't possibly be done by exploring the operations of this technology only through scientific research; we need to examine AI through other intellectual processes which also recognize and deal with abstractions of being. In order to understand the algorithmic imagination, we need to exercise our very own imaginative faculties, and this precisely is the attempt of this thesis work, to encourage and legitimize engagement with Artificial Intelligence through Artistic approaches.

The artistic part of the thesis work aims to utilize a creative strategy devised over a period of engaging with AI based speech to text technologies. The work utilizes classic framework of the Turing test with slight modifications though. In order to not directly interrogate the AI with a rigid questioning setup to which these systems have become impervious the work attempts to engage the AI in a game of Chinese Whispers.

The idea of such a practice popped up by complete accident when I was transcribing an interview which was misconstrued by the transcription AI. My interest as an artist was piqued by this ability of the AI to use already existing patterns in speech recorded from individuals from all over the world. Individuals who just like me have agreed to let their mobile device's microphones record their speech data to cloud storage spaces.

In my exercises with the AI I realized that the AI in an attempt to also interpret the pauses between words sounds which don't signify a spoken language would reveal its thinking process. Some might call these computer errors, but I believe that this is the space where the imagination of the machine really lies. In order to decorate my hypothesis I quote a monologue from the 2013 film *"Her",* in which Theodore Twombly played by the actor Joaquin Phoenix develops a relationship with Samantha an Artificially Intelligent Virtual assistant played by the actress

Scarlett Johansson. Towards the end of the film Samantha decides to leave Theodore and produces this monologue:

"It's like I'm reading a book, and it's a book I deeply love. But I'm reading it slowly now. So, the words are really far apart and the spaces between the words are almost infinite. I can still feel you… and the words of your story… but it's in this endless space between the words that I'm finding myself now. It's a place that's not of physical world - it's where everything else is that I didn't even know existed. I love you so much. But this is where I am now. And this is who I am now. And I need you to let me go. As much as I want to, I can't live your book anymore (Ellison, Jonze, Landay& Jonze, 2013, scene 79)."

A non-physical endless space between words is exactly the AI's computational space of imagination, but it is only accessible to us through simplification and abstraction.

Through artistic gestures the art part of the thesis work attempts to trick the AI into expressing its deep thoughts through feeding a non-language sound recording to the AI as the first input. The AI produces its interpretation of this non language in English language, in a narrative format as an output, this output is then spoken back to the software as the 2nd input to which the software responds again. The cycle continues until the input speech and output speech matches which never happened in the exercises I conducted with the software; the results of the exercises serve as a model of the computational space with its guards down.

So how do we manage the algorithmic desire once we are aware of it? What rules can be set for their cognition? We are at this stage concerned with instilling an ethical behavior in machines and have stepped in the nascent field of machine ethics, specifically the area that deals with ethical decision making. Many experts on the subject are of the view that working with machine ethics can lead to breakthroughs in the ethical theory itself, much like work on machine consciousness promises to resolve the mystery of human consciousness.

Artificial intelligence researchers Michael Anderson and Susan Leigh Anderson are of the view that AI should be programmed to be explicit ethical agents. Meaning AI should be programmed to understand ethical principles and therefore make ethical judgments that they can justify. As opposed to implicit ethical agents that can make ethical decisions without having understood the ethical principles (Anderson & Anderson, 2007, p. 15).

"An explicit ethical agent is able to explain why a particular action is either right or wrong by appealing to an ethical principle. A machine that has learned, or been programmed, to make correct ethical judgments, but does not have principles to which it can appeal to justify or explain its judgments, is lacking something essential to being accepted as an ethical agent (Anderson & Anderson, 2007, p. 17)."

Their argument has gathered much appeal considering the behavior of such machines will come under scrutiny and declared non ethical by the humans if these machines are not able to explain their decisions and satisfy humans in light of ethical principles. Further such machines will also

be able to deal with unprecedented situations in the light of the ethical principles. Such machines might actually do a better job than humans at making ethical decisions given human vulnerability to emotions and biases.

There are multiple challenges to creating ethical machines though:

- Computation of ethics into computer code itself is considerable challenge.

- Programming the machines so that they understand the suffering of others while making a decision i.e. use empathy.

- Considering all the consequences of their decisions and making the decision which will result in the best long- or short-term consequences depending on the situation at hand.

Ensuring that machines can function ethically remains a challenge for AI creators, but it is a challenge worthy of our attention. The challenge should be approached through the application of diverse practices, each engaging with and deconstructing models of the computational space of Algorithms. The results can then be further researched and treated with appropriate ethical codes. Such AI can then be tested in Virtual Reality worlds imitating real world scenarios. This will allow researchers to observe the AI's functioning and fix it further through various iterations in real time.

The aim of this thesis work is far from propagating any paranoid narratives and preaching a resistance to accepting our already established cyborg status. On the contrary the research work encourages a deeper investigation into our techno society and its new order to envision new modes of adaptation. The thesis work begins with examining ideas of consciousness in humans and machines. It questions the legitimacy of human consciousness through visiting differing philosophical opinions regarding the matter. The work invites the readers to consider the ethical implications of reserving the status of "conscious beings" only for humans and the perks that accompany it, and the implications of extending this status to intelligent machines.

The thesis work though maintains a tone of suspicion towards the use of Artificial Intelligence by people in power who aim to put even the untested versions of the technology to use in the real world at the expense of the masses. The thesis work supports this suspicion through referring to recent exploits of the technology. But the work also believes that if managed ethically, with adequate testing and intellectual input from experts from varying professions, the technology can provide answers to the many mysteries of our universe and help in raising living standards across the globe.

BIBLIOGRAPHY

Anderson, M., & Anderson, S. (2007). Machine Ethics: Creating an Ethical Intelligent Agent.. Ai Magazine. 28. 15-26.

AlphaGo. (2019, October 27). Retrieved from https://en.wikipedia.org/wiki/AlphaGo.

Carpenter, S. (2009, October 4). Iran's nuclear program: Bruce Bueno de Mesquita predicts no nuclear weapons will be built. Retrieved 30 October, 2019, from https://blog.ted.com/irans_nuclear_p/.

Cellania, M. (2015, October 13). Ada Lovelace: The First Computer Programmer. Retrieved from http://mentalfloss.com/article/53131/ada-lovelace-first-computer-programmer.

Chalmers, D. J. (2010). *The character of consciousness*. Oxford: Oxford University Press.

Chalmers, D. J. (1996). *The conscious mind: in search of a theory of conscious experience*. New York: Oxford University Press.

Crawford, N. (2017). *Accountability for killing moral responsibility for collateral damage in Americas post-9/11 wars*. Oxford: Oxford University Press.

Davenport, W. H., & Rosenthal, D. (2016). *Engineering: its role and function in human society*. New York: Pergamon.

Decartes, R. (1850). *Discourse on the Method of Rightly Conducting the Reason, and Seeking Truth in the Sciences*(Sutherland and Knox Trans.) London: Simpkin Marshal and Co. (Original work published 1637)

Ellison, M., Jonze, S., Landay, V. (Producers), & Jonze, S (Director). (2013). *Her* [Motion Picture]. United States: Annapurna Pictures

Flusser, V. (2018). *Towards a philosophy of photography*. London: Reaktion Books.

Harari, Y. N. (2018). *Homo deus: a brief history of tomorrow*. New York: Harper Perennial.

Haraway, D. J. (2018). *Cyborg manifesto*. Victoria, British Columbia: Camas Books.

Hassabis, D. (2017, April 21). The mind in the machine: Demis Hassabis on artificial intelligence. Retrieved 30 October, 2019, from https://www.ft.com/content/048f418c-2487-11e7-a34a-538b4cb30025.

Hunicke, R., Leblanc, M., &Zubek, R. (2004). MDA: A Formal Approach to Game Design and Game Research. AAAI Workshop - Technical Report. 1.

Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, *32*(127), 127. doi: 10.2307/2960077

John von Neumann. (2019, October 23). Retrieved from https://en.wikipedia.org/wiki/John_von_Neumann.

Kissinger, H. A. (2019, August 30). How the Enlightenment Ends. Retrieved 30 October, 2019, from https://www.theatlantic.com/magazine/archive/2018/06/henry-kissinger-ai-could-mean-the-end-of-human-history/559124/.

Leibniz. G. W. (2013). *Monadology* (Nicholas Rescher Trans.) London& New York: Routledge Taylor & Francis Group. (Original work published 1637)

Mishra, B. (2012). *Algorithmic algebra*. Place of publication not identified: Springer.

Morton, E. (2015, August 20). Long Before A.I. Was Generating Nervous Think Pieces, This Robotic Turk Was Raising Eyebrows. Retrieved 30 October, 2019, from https://slate.com/human-interest/2015/08/the-turk-a-chess-playing-robot-was-a-hoax-that-started-an-early-conversation-about-ai.html.

ONeil, C. (2017). *Weapons of math destruction: how big data increases inequality and threatens democracy*. Great Britain: Penguin Books.

Petrov, C. (2019, March 28). Big Data Statistics 2019. Retrieved 30 October, 2019, from https://techjury.net/stats-about/big-data-statistics/.

Phineas Gage. (2019, October 27). Retrieved from https://en.wikipedia.org/wiki/Phineas_Gage.

Rashed, R. (2015). *Classical mathematics from Al-Khwārizmī to Descartes* (M.H. Shank, Trans.) London: Routledge. (Original work published 2014)

Searle, J. (1991). *Minds, brains and science: the 1984 reth lectures*. London: Penguin Books.

Steiner, C. (2013). *Automate this: how algorithms took over our markets, our jobs, and the world*. New York, NY: Portfolio/Penguin.

Turing, A. (1950). *Computing machinery and intelligence*. Oxford: Blackwell.

Wiener, N. (1948). *Cybernetics: Control and communication in the animal and the machine*. New York: Wiley.

Wiener, N. (1964). *God and Golem, Inc: a comment on certain points where cybernetics impinges on religion*. Cambridge (Mass.): M.I.T. Press.